

An improved model for the formation times of dark matter haloes

Carlo Giocoli^{1,2}, Jorge Moreno¹, Ravi K. Sheth¹ & Giuseppe Tormen² ^{*}

¹ *University of Pennsylvania, 209 South 33rd Street, Philadelphia, PA 19104 – 6396, USA*

² *Dipartimento di Astronomia, Università degli Studi di Padova, Vicolo dell'osservatorio 2 I-35122 Padova, Italy*

ABSTRACT

A dark matter halo is said to have formed when at least half its mass has been assembled into a single progenitor. With this definition, it is possible to derive a simple but useful analytic estimate of the distribution of halo formation times. The standard estimate of this distribution depends on the shape of the conditional mass function—the distribution of progenitor masses of a halo as a function of time. If the spherical collapse model is used to estimate the progenitor mass function, then the formation times one infers systematically underestimate those seen in numerical simulations of hierarchical gravitational clustering. We provide estimates of halo formation which may be related to an ellipsoidal collapse model. These estimates provide a substantially better description of the simulations. We also provide an alternative derivation of the formation time distribution which is based on the assumption that haloes increase their mass through binary mergers only. Our results are useful for models which relate halo structure to halo formation.

Key words: galaxies: halo - cosmology: theory - dark matter - methods: numerical

1 INTRODUCTION

Large samples of galaxy clusters selected in the optical (Miller et al. 2005; McKay et al. 2005) and other bands (ACT, SPT) will soon be available. These cluster catalogs will be used to constrain cosmological parameters. The tightness of such constraints depends on the accuracy with which the masses of the clusters can be determined from observed properties. These properties are expected to depend on the formation histories of the clusters. Models of cluster formation identify clusters with massive dark matter halos, so understanding cluster formation requires an understanding of dark halo formation.

There is also some interest in using the distribution of galaxy velocity dispersions (Sheth et al. 2003) to constrain cosmological parameters (Newman & Davis 2002). The velocity dispersion of a galaxy is expected to be related to the concentration of the halo which surrounds it, and this concentration is expected to be influenced by the formation history of the halo (Tormen 1998; Bullock et al. 2001; Wechsler et al. 2002). Hence, this program also benefits from understanding the formation histories of dark matter halos.

In hierarchical models, the formation histories of dark matter halos are expected to depend strongly on halo mass—

massive halos are expected to have formed more recently (Press & Schechter 1974). But quantifying this tendency requires a more precise definition of what one means by the ‘formation time’. Lacey & Cole (1993) provided a simple definition—it is the earliest time when a single progenitor halo contains half the final mass. For this definition, they showed how to estimate the distribution of halo formation times. Sheth & Tormen (2004) provide associated expressions for the joint distribution of formation time and the mass at formation. This estimate depends on the distribution of progenitor masses at earlier times, and Lacey & Cole used the assumption that halos form from a spherical collapse to estimate this conditional mass function. However, a model based on ellipsoidal collapse provides a more accurate description of the abundances of dark halos (Sheth, Mo & Tormen 2001) and of their progenitors (Sheth & Tormen 2002). Hence, one expects to find that the ellipsoidal collapse model also provides a better description of halo formation. The primary goal of the present work is to study if this is indeed the case.

Section 2 reviews Lacey & Cole’s argument, and describes how we estimate the ellipsoidal collapse based progenitor mass functions. Explicit formulae are provided in Appendix A. Section 3 compares the model predictions for halo formation with measurements in simulations of hierarchical gravitational clustering. The simulations, and how we estimate halo formation times in them, are described in

^{*} Email: carlo.giocoli@unipd.it, jmoreno@physics.upenn.edu, shethrk@physics.upenn.edu, giuseppe.tormen@unipd.it.

Section 3.1. Sections 3.2 and 3.3 show that the spherical collapse model predicts lower formation redshifts than are seen in the simulations, particularly at low masses. Insertion of ellipsoidal collapse based expressions in Lacey & Cole's formalism provides substantially increased accuracy. However, the agreement is not perfect—although the ellipsoidal collapse based expressions provide a good description of how the median formation redshift decreases with halo mass, it predicts formation time distributions which are broader than seen in the simulations. A final section summarizes our results and discusses possible reasons for the discrepancy.

There is some confusion in the literature about the relation between halo formation as defined by Lacey & Cole (1993), and a quantity which arises in binary merger models of hierarchical clustering. The relation between this other quantity and the one defined by Lacey & Cole is clarified in Appendix B.

2 THE FORMATION TIME DISTRIBUTION

Following Lacey & Cole (1993), we will define the formation time of a halo as the earliest time that at least half of its mass has been assembled into a single progenitor.

2.1 Relation to the progenitor mass function

Consider an ensemble of halos of mass M at time T , and let $N(m, t|M, T)$ denote the average number of progenitors of these halos that have mass $m < M$ at time $t < T$. Since a halo can have at most one progenitor of mass $m > M/2$, the fraction of halos which have a progenitor of mass $m > M/2$ at time t is

$$\int_{M/2}^M dm N(m, t|M, T).$$

But, because they have a progenitor of mass $m > M/2$, these halos are also the ones which formed at some $t_f < t$. Hence,

$$\int_0^{t_f} dt p(t|M, T) = \int_{M/2}^M dm N(m, t_f|M, T), \quad (1)$$

where $p(t|M, T)$ denotes the probability that a halo of mass M at T formed at time t . Differentiating with respect to t_f yields

$$p(t_f|M, T) = \frac{d}{dt_f} \int_{M/2}^M dm N(m, t_f|M, T) \quad (2)$$

$$= \int_{M/2}^M dm \frac{dN(m, t_f|M, T)}{dt_f}. \quad (3)$$

Evidently, the formation time distribution is closely related to the distribution of progenitor masses and its evolution. Different estimates of the progenitor mass function will result in different formation time distributions.

In what follows, we will estimate the distribution of halo formation times using equation (2). Thinking along the lines of equation (3) instead shows how $p(t_f|M, T)$ can be related to quantities which arise naturally in binary merger models of hierarchical clustering. Appendix B provides details.

Lacey & Cole (1993) used estimates for $N(m, t|M, T)$ which were derived from a model in which halos form from

a spherical collapse (Gunn & Gott 1972; Press & Schechter 1974). In this case, they showed that the distribution of formation times could be scaled to a universal form:

$$p(\omega) = 2\omega \operatorname{erfc}(\omega/\sqrt{2}) \quad (4)$$

where

$$\omega = \sqrt{q} \frac{\delta_{\text{sc}}(z_f) - \delta_{\text{sc}}(z_0)}{\sqrt{S(M/2) - S(M)}} \quad (5)$$

and $q = 1$. Here $\delta_{\text{sc}}(z)$ is the overdensity required for spherical collapse at z , and $S(M)$ is the variance in the linear fluctuation field when smoothed with a top-hat filter of scale $R = (3M/4\pi\bar{\rho})^{1/3}$, where $\bar{\rho}$ is the comoving density of the background. In essence, ω is simply a scaled time variable: in an Einstein de-Sitter background cosmology $\omega \propto (z_f - z_0)$, where the constant of proportionality depends on the final mass M .

Strictly speaking, this expression is valid for a white-noise power spectrum ($S(m) \propto m^{-1}$), but it has been found to provide a reasonable approximation for more general power spectra as well. For this distribution, the median value of ω is 0.974, and the associated cumulative distribution is

$$P(>\omega) = \sqrt{\frac{2}{\pi}} \omega e^{-\omega^2/2} + (1 - \omega^2) \operatorname{erfc}\left(\frac{\omega}{\sqrt{2}}\right). \quad (6)$$

However, recent work has shown that the spherical collapse model predicts fewer massive halos and more intermediate mass halos than are seen in simulations of hierarchical gravitational clustering (Sheth & Tormen 1999). Models in which halos form from an ellipsoidal collapse may be more accurate (Sheth, Mo & Tormen 2001). These models also provide a better description of the progenitor mass function (Sheth & Tormen 2002). In what follows, we will show the effects of substituting the ellipsoidal collapse based expressions for $N(m, t|M, T)$ in equation (1). We will also show that simply taking values of q smaller than one will make equation (6) a good fit to the formation redshift associated to the ellipsoidal collapse.

2.2 Progenitor mass functions in the excursion set approach

We will use two approximations for the progenitor mass function. Both forms are derived from casting the ellipsoidal collapse model in the same language used for the spherical model—the excursion set formalism of Bond et al. (1991). In this formalism, halo abundances at a given time are associated with the first crossing distribution by Brownian motion random walks, of a barrier whose height decreases with time, and may in addition depend on how many steps the walk has taken. Figure 1 illustrates: it shows a walk which starts at some initial $S(M)$, and walks to the right, eventually crossing the barrier at some larger value of $S(m) > S(M)$. Here $S(m)$ is the variance in the initial fluctuation field when smoothed with a tophat filter of comoving scale $R = (3m/4\pi\bar{\rho})^{1/3}$, ($\bar{\rho}$ is the comoving background density), and the height of the walk represents the value of the initial density field when smoothed on scale R . In hierarchical models, $S(m)$ is a monotonically decreasing function of mass m . Hereafter we will use δ_0 and δ_1 to indicate the overdensity associated with the spherical collapse model respectively at redshift z_0 and z_1 .

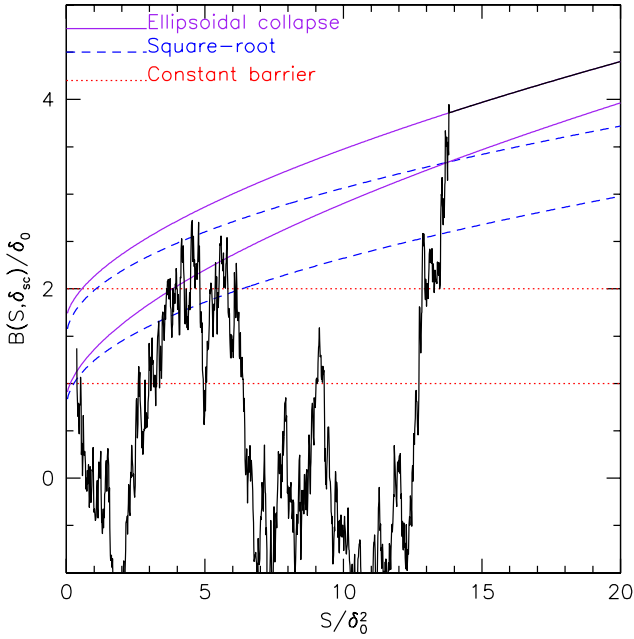


Figure 1. An example of a random walk and the three barrier shapes we consider in this paper. Here δ_0 refers to the critical value associated with spherical collapse overdensity at redshift z_0 . The jagged line is a sample brownian walk absorbed by the barrier associated with ellipsoidal collapse (solid curve). Short-dashed curves show the square-root barriers which we use to approximate the ellipsoidal collapse barrier, and horizontal dotted lines show the constant barrier associated with spherical collapse. The upper set of barriers are associated with collapse at an earlier time.

In the excursion set approach, the shape of the barrier, $B(S)$, depends on the collapse model. If the barrier is crossed on scale $S(m)$, then this indicates that the mass element on which the sphere is centered will be part of a collapsed object of mass m . Bond et al. used the fraction of walks which cross $B(S)$ at $S(m)$ as an estimate of the fraction of mass in halos of mass m : the parent halo mass function. Similarly, the fraction $f(s|S)$ of walks which start from some scale $S(M)$ and height $B(S)$ and first cross the barrier $B(s)$ at some $s > S$, can be used to provide an estimate of the progenitor mass function of M halos:

$$N(m|M) dm = \left(\frac{M}{m}\right) f(s|S) ds \quad (7)$$

(note that $m < M$ because $s > S$). Thus, in the excursion set approach, the shape of $N(m|M)$ depends on how the shape of the barrier changes with mass—it is in this way that the collapse model affects the parent halo and progenitor mass functions.

The barrier shape is particularly simple for the spherical model: $B = \delta_{sc}$ is the same constant for all values of S . Sheth, Mo & Tormen (2001) argued that the barrier associated with ellipsoidal collapse scales approximately as

$$B(S) = \sqrt{q}\delta_{sc} \left\{ 1 + \beta \left[\frac{S(m)}{q\delta_{sc}^2} \right]^\gamma \right\}, \quad (8)$$

where δ_{sc} is the value associated with complete collapse in the spherical model. Following Sheth et al., we set $\beta = 0.5$, $\gamma = 0.6$ and $q = 0.707$; the values of the first two parameters

are motivated by an analysis of the collapse of homogeneous ellipsoids, whereas the value of q comes from requiring that the predicted halo abundances match those seen in simulations. The standard spherical model has $\beta = 0$ and $q = 1$; we will also show the predictions of this model if $\beta = 0$ but $q = 0.707$.

For general γ , exact analytic solutions to the first crossing distribution of barriers given by equation (8) are not available. For this reason, we have studied two analytic approximations which result from addressing the first crossing problem in two different ways. The first approach uses an analytic approximation to the first crossing distribution which Sheth & Tormen (2002) showed was reasonably accurate.

Our second approach is to substitute the ellipsoidal collapse barrier for one which is similar, but for which an exact analytic expression for the first crossing distribution is available. Specifically, when $\gamma = 0.5$, then the barrier height increases with the square-root of S , and the first crossing distribution can be written as a sum of parabolic cylinder functions (Breiman 1966). In this approach, we approximate the progenitor mass function using an expression which would be exact for a square-root barrier, bearing in mind that the square-root barrier is an approximation to the ellipsoidal collapse dynamics. For this barrier shape, we set $\beta = 0.5$ and $q = 0.55$ since these choices result in parent halo and progenitor mass functions which best fit the simulations; Figure 1 illustrates the similarity of the barrier shapes. We will show that both these approximations provide substantially better descriptions of the simulations than does the expression which is based on spherical collapse. Explicit expressions for these barrier crossing distributions are provided in Appendix A. They are related to the progenitor mass function by equation (7).

3 COMPARISON WITH SIMULATIONS

The simulation data analysed below are publicly available at <http://www.mpa-garching.mpg.de/Virgo>. The simulation itself is known as GIF2, and is described in some detail by Gao et al. (2004). It is of a flat Λ CDM universe with parameters $(\Omega_m, \sigma_8, h, \Omega_b h^2) = (0.3, 0.9, 0.7, 0.0196)$. The initial fluctuation spectrum had an index $n = 1$, with transfer function produced by CMBFAST (Seljak & Zaldarriaga 1996).

3.1 The GIF2 simulation

The GIF2 simulation followed the evolution of 400^3 particles in a periodic cube $110 h^{-1} \text{Mpc}$ on a side. The individual particle mass is $1.73 \times 10^9 h^{-1} M_\odot$, allowing us to study the formation histories of haloes which are one order of magnitude smaller in mass than was previously possible. Particle positions and velocities were stored at 50 output times, logarithmically spaced between $1 + z = 20$ and $1 + z = 1$. Halo and subhalo merger trees were constructed from these outputs as described by Tormen et al. (2004). Briefly, halos at a given redshift are identified as spherical objects which enclose the virial density appropriate the cosmological model at that redshift (computed using the spherical collapse model).

A merger history tree was constructed for each halo more massive than $10^{11.5} h^{-1} M_\odot$ (i.e., containing more than

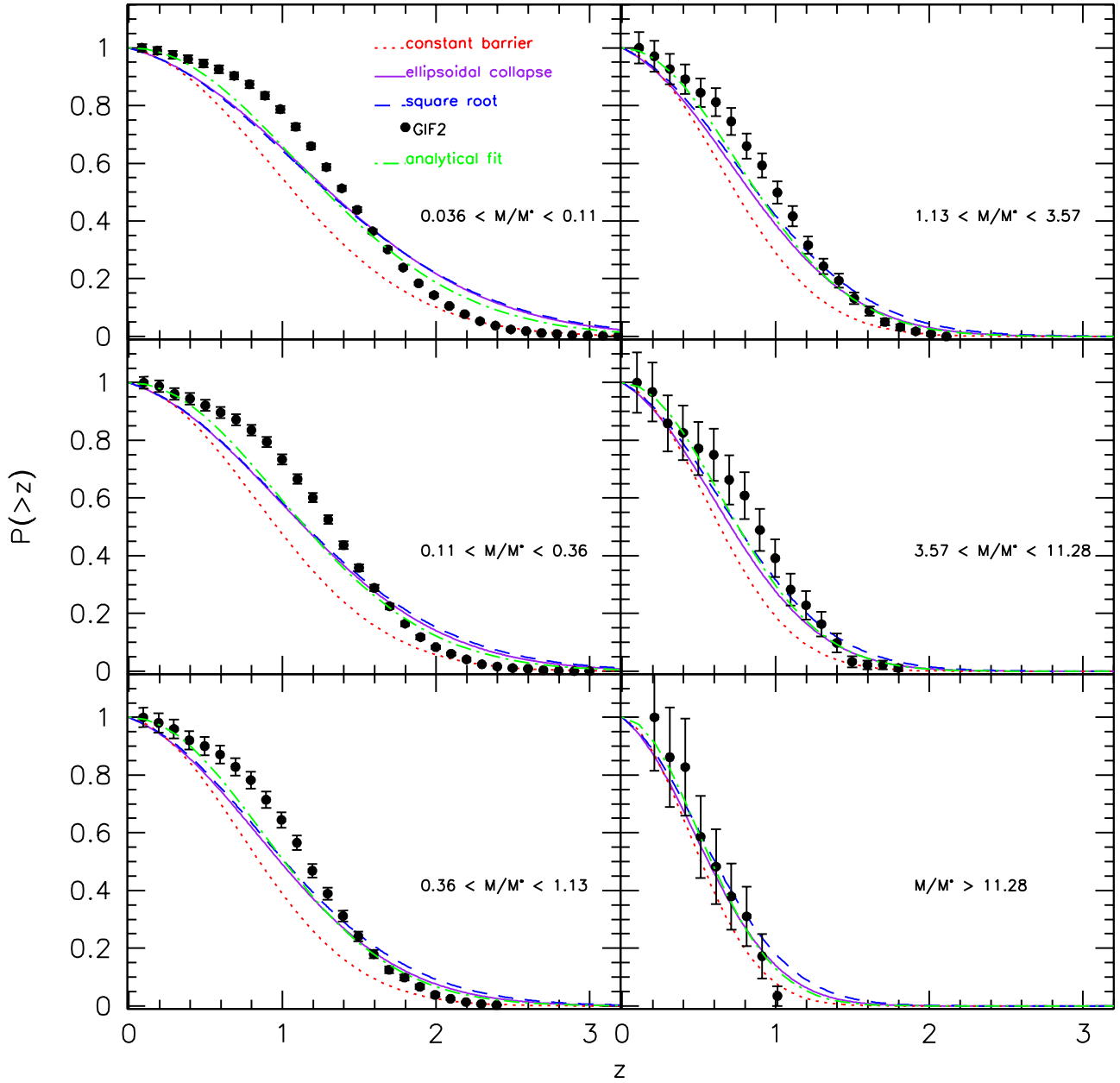


Figure 2. Cumulative distribution of dark halo formation times for halos identified at $z = 0$. From top to bottom, panels show results for halos with masses in the range $\log_{10} M/h^{-1} M_{\odot} : 11.5-12, 12-12.5, 12.5-13, 13-13.5$, and $13.5-14$. Symbols show the measurements in GIF2; dotted curve shows the prediction associated the constant barrier spherical collapse model; dot-dashed curve shows the analytical fit equation (6) with $q = 0.707$. Short-dashed and solid curves show the predictions associated with the square-root and ellipsoidal collapse based models.

180 particles) as follows. The progenitors of a halo of mass M_0 at z_0 are found by identifying in the previous output $z_1 > z_0$ all haloes which contribute any number of particles to M_0 . Of these, the progenitor which contributes the most mass to M_0 is called M_1 . This procedure is then repeated, starting with M_1 at z_1 , considering its progenitors in the previous output time $z_2 > z_1$, and choosing from them the one, M_2 , which contributes the most mass to M_1 . In this way, the

mass of the most massive progenitor is traced backwards in time, to high redshift.

We will present our results as a function of halo mass. We have estimated formation times for 5611 halos with $\log_{10}(M/h^{-1} M_{\odot})$ in the range 11.5 – 12, 2431 in the range 12 – 12.5, 892 in the range 12.5 – 13, 341 in the range 13 – 13.5, 92 in the range 13.5 – 14, and 29 in the range 14 – 14.5. These abundances are well fit by the formula of

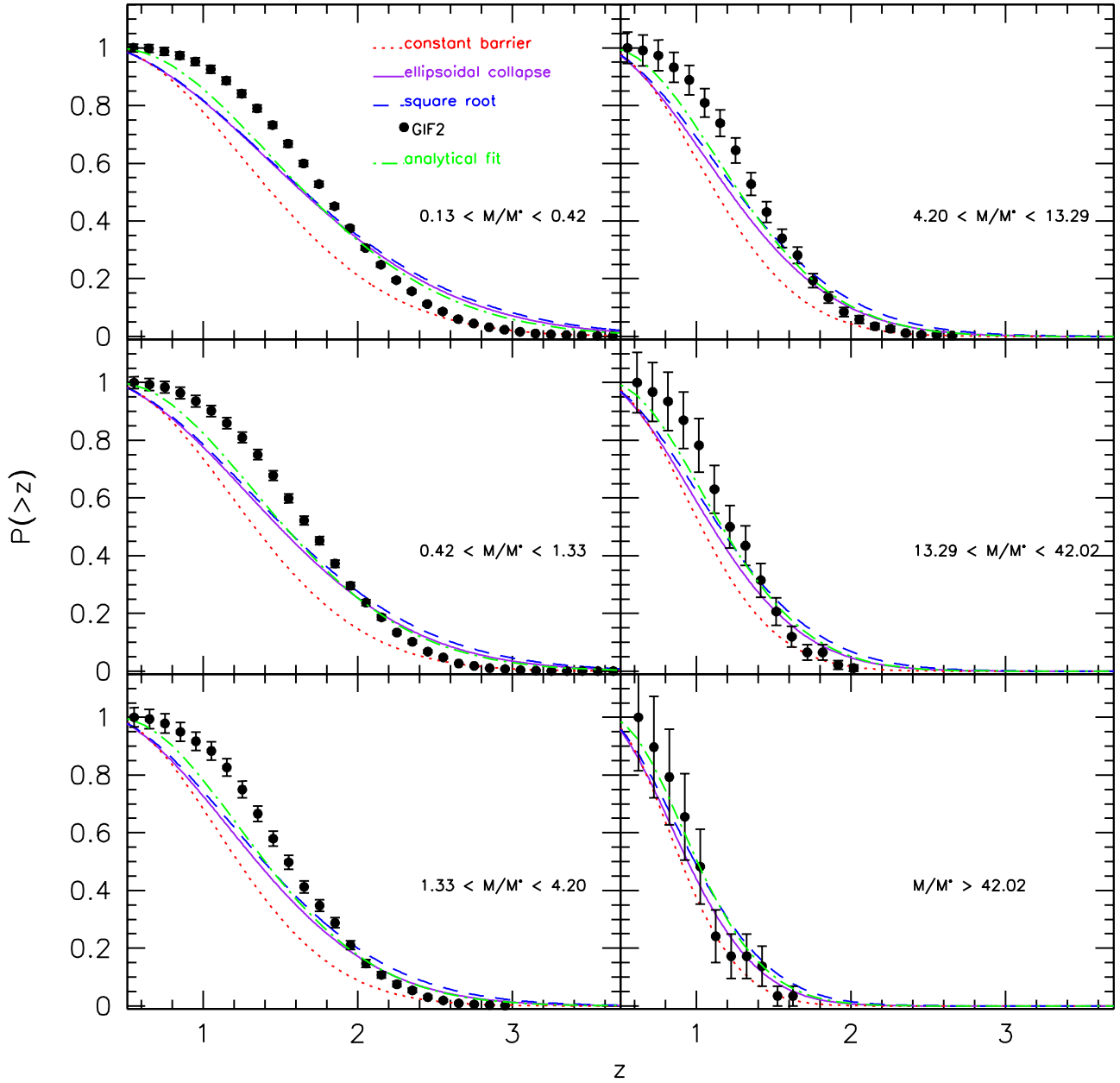


Figure 3. Same as previous figure, but now for halos identified at $z = 0.5$. As before, the spherical model predicts lower formation redshifts than observed. Whereas the ellipsoidal collapse based expressions predict the median formation redshift reasonably well, they predict a broader range of redshifts than are observed.

Sheth & Tormen (1999), so they are in better agreement with the ellipsoidal rather than spherical collapse model.

3.2 Cumulative distribution of formation times

Figure 2 shows the cumulative distribution of dark halo formation redshifts (i.e., equation 1) for halos identified at $z = 0$. In general the analytical expressions obtained from the two different approximate solutions to the ellipsoidal collapse barrier problem—the (Sheth & Tormen 2002) approx-

imation for the first crossing distribution of the ellipsoidal collapse barrier, or the exact expression for the first crossing distribution of the square root barrier—are complicated. However, we have found that the predictions of the two ellipsoidal based models are quite well approximated by the expression for the spherical model with $n = 0$, equation (6), by simply changing the value of q .

Different panels show results for the mass bins described above. The points show measurements in the GIF2 simulation, and the four curves show the formation time distributions associated with the Λ CDM spherical collapse (dot-

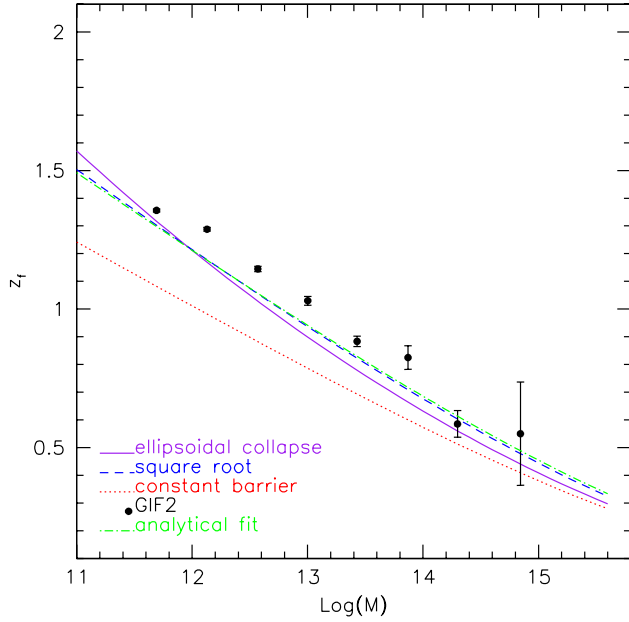


Figure 4. Median formation redshift for halos identified at $z = 0$ as a function of halo mass. Points with (Poisson) error bars show our measurements in the GIF2 simulation. Smooth curves show the median formation times associated with three different models of halo formation: spherical collapse (dotted), ellipsoidal collapse (solid) and the square-root barrier approximation (short-dashed). Dot-dashed curve shows the prediction of equation (10) with $q = 0.707$.

ted), ellipsoidal collapse (solid), square-root barrier approximation (dashed) and with equation (6) using $q = 0.707$ (dot-dashed).

Again, the spherical collapse model severely underestimates the redshifts of halo formation. The two ellipsoidal collapse based estimates fare better, in the sense that they predict median formation redshifts which are closer to those seen in the simulation.

However, both ellipsoidal based estimates clearly predict a wider range of formation redshifts than is seen in the simulation—at fixed mass, the distribution of halo formation redshifts is narrower than predicted. This remains true for the formation time distribution of halos identified at $z = 0.5$ shown in Figure 3.

3.3 Mass-dependence of median formation time

Figure 4 shows the median formation redshift of halos identified at $z = 0$ as a function of halo mass. Points show our measurements in the GIF2 simulation. The simulation shows that massive halos formed more recently: the median formation redshift decreases with halo mass.

Smooth curves show the formation time distributions associated with three different models of halo formation: Λ CDM spherical collapse (dotted), ellipsoidal collapse (solid) and the square-root approximation (short-dashed). In all cases, the model predictions for the median formation redshift were obtained by finding that \bar{z}_f at which

$$\int_{M/2}^M dm N(m, z_f | M, z) = \frac{1}{2}. \quad (9)$$

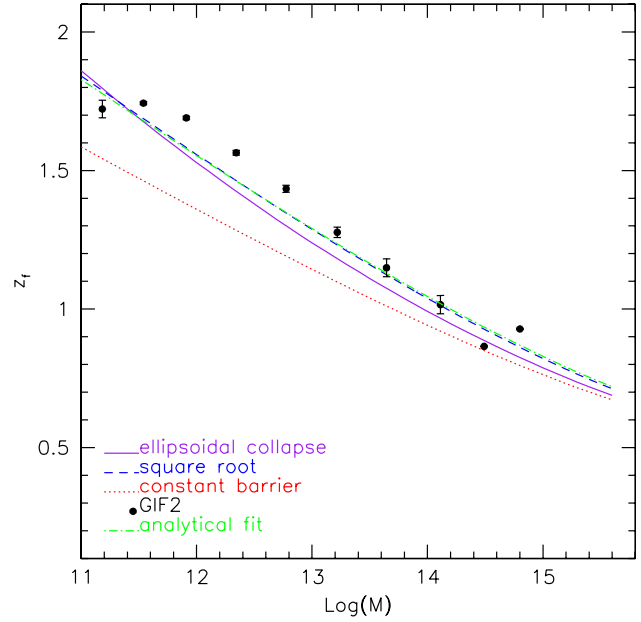


Figure 5. Same as previous Figure, but for halos identified at $z = 0.5$ in the GIF2 simulation.

The Figure shows that halos less massive than $10^{14} h^{-1} M_\odot$ clearly form at higher redshifts than predicted by the spherical collapse model (dotted line). Both our estimates of the ellipsoidal collapse prediction are in substantially better agreement with the simulations. The prediction associated to equation (6) is given implicitly by inserting the median rescaled formation redshift $\bar{\omega} = 0.974$ into equation (5), providing:

$$\delta_{sc}(\bar{z}_f | M, z_0) = \delta_{sc}(z_0) + \frac{0.974}{\sqrt{q}} \sqrt{S(M/2) - S(M)}. \quad (10)$$

The choice $q = 0.707$ is shown by the dot-dashed line in the Figures; setting $q = 0.6$ yields even better agreement with the simulations.

Fitting functions, accurate to a few percent, for $\delta_{sc}(z)$ and $S(m)$ are available in the literature. For instance,

$$\delta_{sc}(z) \approx D_+(z) \frac{3}{20} (12\pi)^{2/3} \Omega_m^{0.0055} \quad (11)$$

(Navarro et al. 1997), where Ω_m is the ratio of the background to the critical density at z , and

$$D_+(z) = \frac{5}{2} \Omega_m \left[\Omega_m^{4/7} - \Omega_\Lambda + \left(1 + \frac{1}{2} \Omega_m \right) \left(1 + \frac{1}{70} \Omega_\Lambda \right) \right]^{-1} \quad (12)$$

(Carroll et al. 1992), and, for a CDM power spectrum,

$$S(m) = A \left(1 + 2.208 \bar{m}^p - 0.7668 \bar{m}^{2p} + 0.7949 \bar{m}^{3p} \right)^{-4/(9p)} \quad (13)$$

(Taruya & Suto 2000), where A is the normalization factor of the linear theory power spectrum at $z = 0$ (so it depends on σ_8), $\bar{m} = m(h\Gamma)^2 / 10^{12} M_\odot$, where Γ is the parameter which describes the shape of the power spectrum (typically $\Gamma \approx \Omega h$), and $p = 0.0873$.

4 DISCUSSION AND CONCLUSIONS

Lacey & Cole (1993) defined the formation time of an object

as the earliest time when at least half its mass was assembled into a single progenitor. We clarified the relation between this definition of formation time, and a quantity which arises in binary merger models of clustering (Appendix B). We have shown that insertion of spherical collapse based expressions in Lacey & Cole's (1993) formalism for halo formation underestimates the redshifts of halo formation; this is consistent with previous work (Lin, Jing & Lin 2003). Ellipsoidal collapse based expressions are a marked improvement: although they result in formation redshift distributions which are broader than seen in simulations (Figures 2 and 3), they predict the median formation redshift quite well (Figure 4 and 5).

The fact that our predicted formation time distributions are broader than those seen in the simulation can be traced back to the fact that the low redshift progenitor mass functions from the excursion set approach are not in particularly good agreement with the simulations. This is similar to the findings of Sheth & Tormen (2002), who noted that some of the discrepancy was almost certainly due to the idealization that the steps in the excursion set walks are uncorrelated. Recent work has shown that there is some correlation between halo formation and environment (Sheth & Tormen 2004; Gao et al. 2005; Harker et al. 2006; Wechsler et al. 2006)—this almost certainly indicates that a model with uncorrelated steps will be unable to provide a better description of the simulations. Nevertheless, the fact that the median formation redshifts are quite well reproduced by our model does represent progress.

Although the exact expressions associated with our models are complicated, we were able to find a useful fitting formula for the median formation redshift, equation (10), which we hope will be useful in studies which relate the formation times of halos to observable quantities.

ACKNOWLEDGMENTS

We thank L. Marian, E. Ricciardelli, F. Biondi and E. Tescari for helpful discussions, and the Aspen Center for Physics for hospitality during the completion of this work, which was supported in part by NSF Grant 0520647 and by CONACyT–Mexico.

REFERENCES

- Bond, J. R., Cole, S., Efstathiou, G., & Kaiser, N. 1991, *ApJ*, 379, 440
 Breiman L., Proc. 5th Berk. Symp. Math. Statist. Prob. 2, 9
 Bullock, J. S., Kravtsov, A. V., & Weinberg, D. H. 2001, *Apj*, 548, 33
 Carroll, S. M., Press, W. H., & Turner, E. L. 1992, *ARA&A*, 30, 499
 Eke, V. R., Cole, S., & Frenk, C. S. 1996, *MNRAS*, 282, 263
 Gao, L., White, S. D. M., Jenkins, A., Stoehr, F., & Springel, V. 2004, *MNRAS*, 355, 819
 Gao, L., Springel, V., & White, S. D. M. 2005, *MNRAS*, 363, L66
 Gunn, J. E., & Gott, J. R. I. 1972, *Apj*, 176, 1

- Harker, G., Cole, S., Helly, J., Frenk, C., & Jenkins, A. 2006, *MNRAS*, 367, 1039
 Kauffmann, G., Colberg, J. M., Diaferio, A., & White, S. D. M. 1999, *MNRAS*, 303, 188
 Lacey, C., & Cole, S. 1993, *MNRAS*, 262, 627
 Mahmood, A., & Rajesh, R. 2005, *ArXiv Astrophysics e-prints*, arXiv:astro-ph/0502513
 McKay, T. A., et al. 2005, *American Astronomical Society Meeting Abstracts*, 206, #10.12
 Miller, C. J., et al. 2005, *AJ*, 130, 968
 Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, *ApJ*, 490, 493
 Newman, J. A., & Davis, M. 2002, *ApJ*, 564, 567
 Percival, W., & Miller, L. 1999, *MNRAS*, 309, 823
 Press, W. H., & Schechter, P. 1974, *ApJ*, 187, 425
 Seljak, U., & Zaldarriaga, M. 1996, *ApJ*, 469, 437
 Sheth, R. K., & Pitman, J. 1997, *MNRAS*, 289, 66 1057
 Sheth, R. K. 1998, *MNRAS*, 300,
 Sheth, R. K., & Tormen, G. 1999, *MNRAS*, 308, 119
 Sheth, R. K., Mo, H. J., & Tormen, G. 2001, *MNRAS*, 323, 1
 Sheth, R. K., & Tormen, G. 2002, *MNRAS*, 329, 61
 Sheth, R. K., et al. 2003, *ApJ*, 594, 225
 Sheth, R. K. 2003, *MNRAS*, 345, 1200
 Sheth, R. K., & Tormen, G. 2004, *MNRAS*, 349, 1464
 Smoluchowski, M. V. 1916, *Zeitschrift fur Physik*, 17, 557
 Taruya, A., & Suto, Y. 2000, *ApJ*, 542, 559
 Tormen, G. 1998, *MNRAS*, 297, 648
 Tormen, G., Moscardini, L., & Yoshida, N. 2004, *MNRAS*, 350, 1397
 van den Bosch, F. C. 2002, *MNRAS*, 331, 98
 Verde, L., Kamionkowski, M., Mohr, J. J., & Benson, A. J. 2001, *MNRAS*, 321, L7
 Wechsler, R. H., Bullock, J. S., Primack, J. R., Kravtsov, A. V., & Dekel, A. 2002, *ApJ*, 568, 52
 Wechsler, R. H., Zentner, A. R., Bullock, J. S., & Kravtsov, A. V. 2005, *ArXiv Astrophysics e-prints*, arXiv:astro-ph/0512416

APPENDIX A: EXPLICIT EXPRESSIONS FOR BARRIER CROSSING DISTRIBUTIONS

We are interested in walks which start from one barrier $B_0(S)$ and walk to another, $B_1(s)$, with $s > S$. Thus, the barrier to be crossed is of the form

$$B_1(s) - B_0(S) = \sqrt{q}\delta_1 \left[1 + \beta \left(\frac{s}{q\delta_1^2} \right)^\gamma \right] - \sqrt{q}\delta_0 \left[1 + \beta \left(\frac{S}{q\delta_0^2} \right)^\gamma \right]. \quad (\text{A1})$$

In the formulae δ_0 and δ_1 refer to the spherical collapse overdensity at redshift z_0 and z_1 . For barriers of this form, Sheth & Tormen (2002) showed that

$$f(s|S)ds \approx \frac{ds}{s-S} \frac{|T(s|S)|}{\sqrt{2\pi(s-S)}} \times \exp \left\{ - \frac{[B_1(s) - B_0(S)]^2}{2(s-S)} \right\}, \quad (\text{A2})$$

where

$$T(s|S) = \sum_{n=0}^5 \frac{(S-s)^n}{n!} \frac{\partial^n [B_1(s) - B_0(S)]}{\partial s^n}. \quad (\text{A3})$$

When $\beta = 0$ and $q = 1$, this reduces to the spherical collapse based expression used by Lacey & Cole (1993):

$$f(s|S)ds = \frac{ds}{s-S} \frac{\delta_1 - \delta_0}{\sqrt{2\pi(s-S)}} \exp\left\{-\frac{(\delta_1 - \delta_0)^2}{2(s-S)}\right\}. \quad (\text{A4})$$

When $\gamma = 1/2$, then

$$B_1(s) - B_0(S) = \sqrt{q}(\delta_1 - \delta_0) - \beta\sqrt{S} + \beta\sqrt{s}, \quad (\text{A5})$$

and the exact solution to this square-root barrier problem is

$$f(s|S)ds = \exp\left(\frac{w_\beta^2 - \beta^2}{4}\right) \times \sum_\lambda \left(\frac{S}{s}\right)^{\frac{\lambda}{2}} \frac{D_\lambda(w_\beta) D'_\lambda(-\beta)}{I_\lambda(-\beta)} \frac{ds}{2s} \quad (\text{A6})$$

(Breiman 1966; Mahmood & Rajesh 2005), where

$$w_\beta \equiv \frac{\sqrt{q}(\delta_1 - \delta_0)}{\sqrt{S}} - \beta, \quad I_\lambda(-\beta) \equiv \int_{-\beta}^{\infty} dx D_\lambda^2(x),$$

and $D_\lambda(x)$ are parabolic cylinder functions which satisfy

$$D'_\lambda(x) = (x/2) D_\lambda(x) - D_{\lambda+1}(x).$$

The sum is over the roots λ defined by

$$D_\lambda(-\beta) = 0, \quad \text{so} \quad D'_\lambda(-\beta) = -D_{\lambda+1}(-\beta).$$

APPENDIX B: ON THE DIFFERENCE BETWEEN HALO FORMATION AND CREATION

The main text deals with halo formation as defined by Lacey & Cole (1993). Before continuing, it is worth mentioning that there is another context in which the term ‘halo formation’ arises—this is when the halo population is modeled as arising from a binary coagulation process of the type first described by Smoluchowski (1916). In this description, the time derivative of the halo mass function is thought of as the difference of two terms: one represents an increase in the number of halos of mass m from the merger of two less massive objects, and the other is the decrease in the number of m -haloes which results as m -haloes themselves merge with other halos, creating more massive halos as a result:

$$\frac{dn(m,t)}{dt} = C(m,t) - D(m,t) \quad (\text{B1})$$

The gain term, the first on the right hand side above, is sometimes called the halo formation (Sheth & Pitman 1997) or creation (Percival & Miller 1999; Sheth 2003) term. In what follows, we will use the word ‘creation’ to mean this term, and ‘formation’ to mean the quantity studied by Lacey & Cole.

Halo formation and creation are *very* different quantities, as we show below. Nevertheless, they are sometimes used interchangeably in the literature (e.g. Verde et al. (2001)). Here we show explicitly how to compute the Lacey-Cole formation time distribution from the Smoluchowski-type creation and destruction terms, with the primary aim of insuring that the error of confusing one for the other is not repeated.

In the Lacey-Cole picture, the formation time distribution of M -halos identified at T is given by equation (2).

In this case, one first integrates over the conditional mass function, and then takes a time derivative. Here, we will instead take the time derivative inside the integral over mass and compute it before integrating over mass, as suggested by equation (3). In this case, the integrand has the form of a time derivative which plays a central role in the Smoluchowski picture. In particular, we can write

$$\frac{dN(m,t|M,T)}{dt} = C(m,t|M,T) - D(m,t|M,T), \quad (\text{B2})$$

where C and D are now the creation and destruction terms associated with the progenitor mass function, evaluated explicitly in Sheth (2003).

This simple step shows explicitly that $p(t|M,T)$ can be written in terms of Smoluchowski-like quantities:

$$p(t|M,T) = \int_{M/2}^M dm C(m,t|M,T) - \int_{M/2}^M dm D(m,t|M,T). \quad (\text{B3})$$

A little thought shows why this works out so easily. Simply integrating the halo creation rate $C(m,t|M,T)$ over the range $M/2 \leq m \leq M$ overestimates halo formation, since some of the halos created at t , with mass $3M/4 \leq m \leq M$ say, may actually have been created by binary mergers in which one of the halos had a mass in the range $M/2 \leq m \leq 3M/4$. Such creations should not be counted towards halo formation, since formation refers to the earliest time that m exceeds $M/2$. However, it is precisely this double-counting which the second term, the integral of $D(m,t|M,T)$ over the range $M/2 \leq m \leq M$, removes. Note in particular that, whereas the formation time distribution $p(t|M,T)$ is related to Smoluchowski-type quantities, it is *not* the same as $C(m,t)$. For completeness, note that

$$C(m,t) = \int_m^\infty dM n(M,T) C(m,t|M,T), \quad (\text{B4})$$

where $n(M,T)$ denotes the number density of halos of mass M at time T .